



## Ürünler için sahte yorum yazanların etkin olarak belirlenmesi

**Buket Kaya\***

Fırat Üniversitesi, EOSB Elektronik ve Otomasyon Bölümü, Elazığ

[bkaya@firat.edu.tr](mailto:bkaya@firat.edu.tr) ORCID: 0000-0002-6917-5170, Tel: (424) 237 00 00 (8219)

Geliş: 11.05.2018, Kabul Tarihi: 12.06.2018

### Öz

Çevirim içi alışveriş sitelerinin popülerliğinin son yıllarda artması ile birlikte internet kullanıcıları daha fazla sayıda ve miktarda ürünü internet üzerinden satın almaktadır. Bugün Türkiye'nin büyük birçok şehrinde market alışverişi bile internette yapılabilmektedir. İnternet üzerinden alışveriş yapan kullanıcıların, ürünler hakkında diğer kullanıcıların yorumlarına önem verdiği ve çoğu zaman satın alma eğilimlerinin diğer kullanıcıların görüşleri ile birlikte değiştiği birçok çalışma ile tespit edilmiştir. Şirketler kendi ürünleri hakkında olumlu kanı oluşturmak veya rakip ürünler hakkında var olan iyi izlenimi kötüleştirmek yani yorum sistemini manipüle etmek için kimi zaman gönüllü kimi zamanda paralı yorumcular tutabilmektedir. Böylece objektiflikten uzak ve taraflı yorum yazdırarak satın alacak kullanıcıların fikirlerini değiştirmeye çalışmışlardır. Bu yorumları insan eli ile düzeltmek mümkün olmayacak kadar çoktur. Bundan dolayı sahte yorum ve sahtekarları tespit etmek için birçok yaklaşım sunulmuştur. Bu yaklaşımlar ile sahte yorumlar ve sahtekarlar tespit edilerek yorum sistemlerinin güvenilirliği sağlanabildiği gibi kullanıcıların manipüle edilmemiş yorum okumaları sağlanarak yorumlara olan güven arttırabilir. Bu çalışmada sahte yorum yazarları bulmak için önerilen dört farklı metot entegre edilerek yeni bir yaklaşım sunulmuştur. Amazon ürün yorumları üzerinde test edilen yöntem, daha doğru sonuçlar bulduğunu göstermiştir.

**Anahtar Kelimeler:** Sahte Yorum; Sahte Yorumcu; Makine Öğrenmesi; Sınıflandırma;

\* Yazışmaların yapılacağı yazar

## Giriş

İnternet üzerinden yapılan alışveriş sayısının ve kapsamının her geçen gün artması ile birlikte milyonlarca ürün ve hizmet çevrimiçi olarak satın alınabilmektedir. Çok fazla alternatif seçenek olduğu için kullanıcıların beklentilerini karşılayan en iyi ürünü bulmak zor bir süreçtir. Bir ürün veya hizmeti seçmek için en güvenilir yaklaşımlardan biri, onları daha önce denemiş olan kullanıcıların deneyimini dinlemektir. Forum veya çevrimiçi pazar sitelerindeki yorum sayfaları, bireylerin ürün ve hizmetler hakkındaki deneyimlerini paylaştığı platformlardır. Bireyler tarafsız ve profesyonel yorumlar yazdığı gibi önyargılı olabilecek yorumlarda yazabilir ve/veya okuyabilir. Bununla birlikte şirketler, mallarını veya hizmetlerini iyileştirmek için yorum sistemlerinden yararlanarak görüş toplayabilir ve rakiplerini izleyebilirler. Ancak, yorum sistemlerinin popüleritesinin yüksek olmasından dolayı bazı kişi ve/veya kurumlar sahte yorumlarıyla bazı ürünlerin tanıtımını yapmaya çalışabilirler. Tanıtım amaçlı yapılan bu yorumların tespiti ve kaldırılması, potansiyel müşterilerin yanıltılmasının önlenmesi ve pazarın etik dışı durumları ortadan kaldırması için önem arz etmektedir (Crawford vd., 2015).

Sahte yorumları tespit etme çok iyi çalışılmamış nispeten yeni bir araştırma problemidir. Bu probleme farklı araştırmacılar değişik boyutlardan bakmışlardır. İlk çalışmalar Jindal ve Liu (Jindal vd., 2007; Jindal vd., 2008) ve Thet ve diğ. (Thet vd., 2010) tarafından yapılmıştır. Tavsiye edici sistemlerle birlikte E-posta, ve Web sahtekarlıkların kapsamlı bir şekilde incelendiği ancak sahte yorumlarla ilgili çok az çalışma yapıldığını iddia etmişlerdir. Lim ve arkadaşları (Lim vd., 2010) sahte yorumları bulmak için yorum yapanın davranışlarına odaklanmanın daha iyi bir fikir olduğunu belirtmişlerdir. Bu yazarlar, yorum yapanların davranışlarını modellediği iki önemli görüşü tanımladılar.

Bunlardan birincisi, sahte yorum yazıcılar etkilerini en üst düzeye çıkarmak için belirli ürün

veya ürün gruplarını hedefleyebilirler. Diğeri ise, sahte yorum yazıcılar ürün değerlendirmelerde diğerk yorumculardan sapma gösterirler. Lim ve arkadaşları (Lim vd., 2010) ayrıca sahte yorumları bulmada; sahte yorumcunun davranışından kazanılabilecek bilgi miktarının, yorumlardan kazanılacak bilgi miktarından çok daha fazla olduğundan yorumlardan ziyade yorumculara yoğunlaşmanın daha iyi bir fikir olduğunu vurgulamışlardır.

Diğerk bir çalışmada Mukherjee ve diğ. (Mukherjee vd., 2011) sahte yorum gruplarını yani ürünleri teşvik edici veya kötüleyici yorumları birlikte yazan grupları bulmaya çalıştılar. Bu yazarlar aşağıdaki senaryolardan herhangi birinin oluşması durumunda aynı ürüne sürekli olarak yorum yazan bir yorumcu grubun olası bir sahte yorum yazıcı olabileceğini kabul etmiştir: (1) kısa bir sürede yorum yazanlar, (2) bir tek ürün için yorumların oranı, o ürün için geri kalan oranlardan sapma gösterenler, (3) yorumların içeriği birbirinin aynı olanlar, (4) ürün eleştirilmeye açık olmaya başladıktan hemen sonra yorum yazanlar, (5) bir ürünü yorumlayanların büyük bir bölümünü oluşturanlar, (6) ürünlerin önemli miktarı üzerinde birlikte çalışanlar.

Jindal ve arkadaşları (Jindal vd., 2010) ise yorumculardan olağandışı davranışı ayırarak sahte yorumların belirlenebileceğini iddia etmiştir. Böylece, birliktelik kural madenciliğinden faydalanan veriden beklenmedik kurallar keşfedilerek bu tip problemi adresleyen alandan bağımsız bir çerçeve tasarlanabilir.

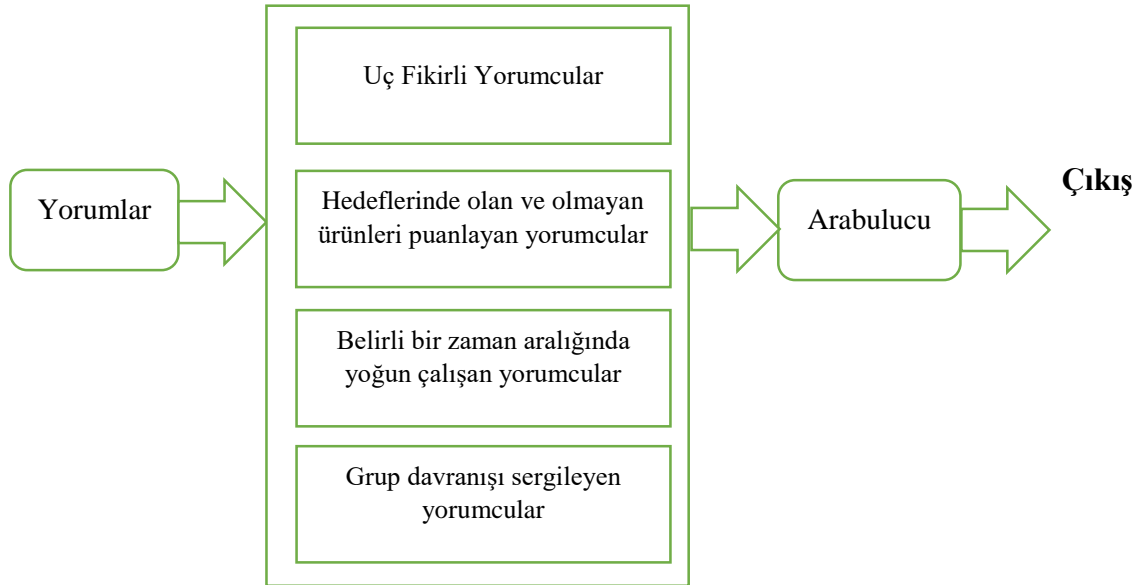
Bu çalışmanın amacı, ürünlerle ilgili deneyimlerini anlatan gerçek alıcıların gerçek düşüncelerini yansıtmak için yazılan görüşleri bulmaktır. Bununla birlikte, şirketler için belli bir menfaat ve para karşılığı yorumlar yazanları ve bu şekilde çeşitli hileler uygulayarak sıradan yorumları gibi görünen ama ücretli yorumlar yayınlayanları yakalamak ve sistemden çıkarmaktır. Aslında bu tür yorumları bulmak çok karmaşık bir iştir ve özellikle bu

yorumları yapanlar, görevlerinde çok profesyonel olduklarında işlem daha da karmaşılaşır (Lim vd., 2010; Heydari vd., 2015). Ayrıca bu işin kullanıcılar veya görevliler tarafından tek tek elle yapılması da düşünülemez. Sonuç olarak, bu gerçekçi olmayan yayınları saptamak için otomatik bir sistem gereklidir. Bu sistem, yanıltıcı fikirleri belirleme yönünde çeşitli teknikler kullanılmalıdır, diğer bir deyişle, profesyonel olarak yazılan ama gerçek olmayan yorumları bile algılayabilmek için sorunu farklı bakış açılarından düşünmelidir. Birden çok yöntemin entegre edilerek hibrit bir sistem olarak sunmanın nedeni, sahte yorum gönderenlerin gözden geçirme sistemlerini yanıltmaya yönelik farklı yaklaşımları kullanması ve bu nedenle farklı şekillerde ele alınması gerektiğidir. Bundan dolayı, gerçekçi olmayan mesajları bulma sorununun çözümü için tüm yorumları ve olası tüm yaklaşımları hesaba katmalıdır. Bu araştırmada, yukarıda açıklanan sorunu çözmek için bazı yaklaşımlar önerilmiştir. Her yaklaşım

problemi farklı bir perspektiften ele alır. Bu yaklaşımlar, anormal eğilimler gösteren, nadir veya şüpheli davranışlar sergileyen yorumcuları bulma ve grup etkinliklerini araştırma gibi çeşitli stratejilere dayalıdır .

## Önerilen Yöntem

Bu bölümde önerilen sistem tanıtılacaktır. Sistemde birbirinden farklı 4 sahte yorumcu bulma yaklaşımı bir araya getirilmiştir. Her bir yaklaşımın sahte yorumcu bulma yöntemi farklıdır. Önerilen yöntem bu sayede nadir ve şüpheli davranışlar sergileyen ya da grup etkinliği gösteren sahte yorumcuları birlikte bulabilir. Şekil 1 önerilen sistemin mimarisini gösterir. Bu mimariye göre ürünler için toplanan yorumlar, sahte yorumcu kısmına gelir. Bu kısımda 4 farklı sahte yorumcu bulma yaklaşımı vardır.



Şekil 1. Önerilen sistemin mimarisini

İlk yaklaşım uç fikirli yorumcuları bulur. Bu yaklaşıma göre, insanlar çoğunlukla hedeflerindeki bir ürün için ya tamamen tatminkar ya da tamamen şikayetkar olmazlar (Lim vd., 2010). Bir ürünün mutlaka olumsuz bir yönü ya da tam tersi bir olumlu yönü vardır. En

basit sahte yorumcu bulma yöntemidir. Bu modülde yapılan belirli bir sahte yorumcunun, yorumlarının büyük bir çoğunluğu-ki uygulama sonuçlarında bu değer %90 olarak seçilmiştir- en düşük ya da en yüksek değerle oranlanmışsa bu

yorumcunun sahte bir yorumcu olma ihtimali çok yüksektir.

İkinci yaklaşımda sahte yorumcular hedeflerindeki ürünü farklı, hedeflerinde olmayan ürünleri ise daha farklı oylarlar (Lam vd., 2004). Öyle ki, bu yöntemde her bir yorumcunun yorumları iki kategoriye ayrılır. İlk kategoride yorumcunun hedefindeki ürün uç bir değerle oranlanır. Diğer kategoride ise, yorumcunun hedefinde olmayan ürünler vardır. Bu ürünleri ise, daha önce o ürün için yapılan yorumların ortalama bir değeri ile oranlar. Bu sayede aslında uç fikirli bir yorumcu olmadığı kanaatini uyandırarak izini kaybetmeye çalışır.

Üçüncü yaklaşımda, belirli bir zaman aralığında yoğun bir şekilde yorum yapan yorumcular hedeflenmiştir (Xie vd., 2012). Bu yöntemde her bir yorumcunun yorumları zaman eksenine göre dizilmiştir. Belirli bir zaman aralığı boyunca pozitif veya negatif yönde birçok yorum yazan yorumcunun sahte yorumcu olma ihtimali yüksektir. Burada önemli bir nokta bu zaman aralığının uygun bir şekilde belirlenebilmesidir. Aralık çok kısa seçilirse, sahte yorumcu bulma olasılığı düşer. Çok büyük seçilirse, sahte yorumcu bulma olasılığı düşer. Çok büyük seçilirse, bu oran da artar. Süreç deneme yanılma yöntemiyle sahte yorumcuların iyi bir sayısı elde edilene kadar devam edilmelidir.

Dördüncü ve son yaklaşımda ise kendi başından ziyade grup halinde çalışan sahte yorumcuların tespit edilmesine çalışılmıştır (Mukherjee vd., 2011). Bu yöntemde öncelikle her bir ürün için, en az üç yorumun olması gereklidir. Bunu yapmasındaki neden, her bir yorumun diğer yorumların ortalama oranlarına olan uzaklıklarının bulunmasıdır. Böylece üç yorumdan sapma gösteren yorum kolaylıkla görülebilir. Bütün ürünler için, bu süreç tekrarlanır. Sonuç olarak bir yorumcu diğer yorumculardan anormallik gösterirse, o yorumcunun sahte yorumcu olma ihtimali yüksektir.

Mimarinin üçüncü kısmı arabulucu olarak adlandırılır. Bu kısımda bulunan sahte

yorumcular aşağıdaki formüle göre büyükten küçüğe doğru sıralanırlar.

$$\text{Öncelik} = \frac{W_i}{SY}, i = 1, 2, 3, 4 \quad (1)$$

Burada  $W_i$ , sahte yorumcuların  $i$  yaklaşım tarafından bulunduğunu belirtir.  $SY$  ise, ilgili yaklaşımlar tarafından bulunan sahte yorumcu sayısıdır. Örneğin, mimariye verilen yorumlardan 20 tane sahte yorumcu 4 farklı yaklaşım tarafından da bulunmuşsa;

$\text{Öncelik}(1,2,3,4) = 4/20$ 'dir ve bu 20 yorumcunun sahte yorumcu olma ihtimali çok yüksektir. Çünkü dört farklı yöntemin sahte yorumcu bulma özelliğini aynı anda taşır. Diğer bir örnek olarak, 50 sahte yorumcu 1., 2., ve 3.yaklaşım tarafından, 60 sahte yorumcu da 2.,3., ve 4. yaklaşımlar tarafından bulunmuşsa,  $\text{Öncelik}(1,2,3) = 3/50$  değeri,  $\text{Öncelik}(2,3,4) = 3/60$ 'dan daha büyük olduğundan, 1,2, ve 3 nolu yaklaşımların bulunduğu 50 yorumcunun sahte olma ihtimali daha yüksektir. Çünkü az sayıda bulunan yorumcunun sahte olma ihtimali daha fazladır. Bu şekilde öncelik sıralamasına göre, yaklaşımlar sıralanır. Bu yöntemde göre, tek bir yaklaşım tarafından en fazla sayıda bulunan yorumcuların sahte olma ihtimali diğerlerine göre daha düşüktür.

## Uygulama Sonuçları

Bu bölümde önerilen yöntemin başarımını test etmek için bazı uygulamalar yapılmıştır. Veri olarak, Amazon ürünleri için yazılan yorumlar kullanılmıştır (He vd., 2016). Yaklaşımlar arasındaki öncelik sıralamasını elde etmek için kullanılan veri kümesinde 1.689.188 yorum, 382.176 ürün ve 143.615 yorumcu vardır. Her bir ürün hakkında yorum yapıldıktan sonra ürün 1 ile 5 arasında bir değerle oylanır. Bu veri tabanında her bir ürün ve her bir yorumcunun en az 5 yorumu bulunmaktadır. Eğitim verisi niteliğindeki bu veri kümesinin önerilen sistem ile elde edilen öncelik sırası ve bulunan yorumcu sayısı Tablo 1'de verilmektedir. Tabloda gösterildiği gibi önerile sistemin amacına uygun olarak sadece iki veya daha fazla yaklaşım tarafından bulunan sahte yorumcular sıralamaya

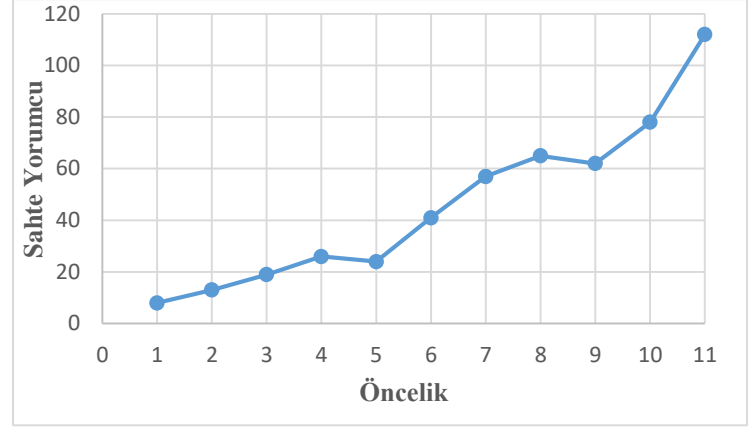
sokulmuştur. Birinci öncelik dört yaklaşımın da bulunduğu 41 sahte yorumcudur. Bu yorumcuların sahte olma olasılığı çok fazladır ve bir sonraki uygulamada önerilen sistemin başarımı daha önceden sahte olduğu bilinen yorumcularla test edilecektir. Üçlü yaklaşımlar içerisinde önceliği en yüksek olan ise 1., 3., ve 4. yaklaşımların birlikte bulunduğu 186 potansiyel sahte yorumcudur. En düşük önceliği ise 2. ve 4. yaklaşımların bulunduğu 13.723 olası muhtemel sahte yorumcudur.

**Tablo 1.** Yaklaşımların öncelik değerleri

Yaklaşım	Bulunan Sahte Yorumcu Sayısı	Öncelik
1, 2, 3, 4	41	1
1, 3, 4	186	2
1, 2, 3	373	3
2, 3, 4	517	4
1, 2, 4	1328	5
1, 3	1860	6
2, 3	3245	7
3, 4	5619	8
1, 2	9755	9
1, 4	12080	10
2, 4	13723	11

Uygulamaların ikinci kısmında önerilen sistemi test etmek için 750'si sahte yorumcu olmayan ve 250'si de sahte yorumcu olan sentetik bir veri kümesi oluşturulmuştur. Bu 250 sahte yorumcu eşit olasılıkla en az iki yaklaşımla bulunacak şekilde üretilmiştir. 11 önceliğe göre bulunan sahte yorum sayısı Şekil 2'de verilmiştir. Buna göre 9 öncelikli yaklaşımın bulunduğu sahte yorumcu sayısı daha önceki uygulamanın sonuçlarına benzer iken sadece 5 ve 9 öncelik numaralı yaklaşımların öncelik sırası değişmiştir.

Bu bölümde yapılan son uygulama, en az iki yaklaşımlı çözümlerin doğruluk değerini elde etme üzerinedir. Tablo 2 bu yaklaşımların doğruluk değerini gösterir. Beklenildiği gibi en yüksek doğruluk dört yaklaşımın da bulunduğu sahte yorumcudur. En düşük doğruluk değerine sahip yaklaşım ise 2 ve 4 numaralı yaklaşımların elde ettiği çözümlerdir.



**Şekil 2.** Eğitim verisindeki önceliğe göre bulunan test verisindeki sahte yorumcu sayısı

**Tablo 2.** Test verisine göre yaklaşımların Doğruluk Değerleri

Yaklaşım	Doğruluk
1, 2, 3, 4	0,83
1, 3, 4	0,82
1, 2, 3	0,81
2, 3, 4	0,81
1, 2, 4	0,79
1, 3	0,78
2, 3	0,78
3, 4	0,77
1, 2	0,70
1, 4	0,69
2, 4	0,65

## Sonuçlar

Bu çalışmanın amacı, ürünlerle ilgili deneyimlerini anlatan gerçek yorumcuların görüşlerini bulmak ve belirli bir menfaat karşılığı yorum yazanların ise yorumlarını çıkarmaktır. Sahte yorumcuları bulmak aslında karmaşık işlerdir. Bu yorumları yazanlar çok profesyonel olduklarında işlem daha da karmaşıklaşır. Ayrıca böyle bir sürecin kullanıcılar tarafından tek tek elle yapılması büyük veri kümeleri için olanaksızdır. Bu çalışmada gerçekçi olmayan oylamalar yapan sahte yorumcuların tespiti için bir sistem önerilmiştir. Sistem yanıltıcı fikirleri belirleme yönünde dört farklı yaklaşım kullanılmaktadır. Birden fazla yaklaşımın kullanılma nedeni, sahte yorum gönderenlerin yanıltmaya yönelik değişik yöntemleri

kullanmasıdır. Yaklaşımlar genel olarak bireysel veya grupsal davranışları göz önüne alır. Amazon ürünlerin yorumları üzerinde test edilen sistem, yüksek doğruluk değerleriyle etkin çözümler sunmaktadır.

## Kaynaklar

- Crawford, M., Khoshgoftaar, T. M., Prusa, J.D., Richter A.N., Al-Najada, H., (2015). Survey of Review Spam Detection using Machine Learning Techniques, *Journal of Big Data*, 2:23.
- He, R., McAuley, J., (2016). Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering, *WWW '16 Proceedings of the 25th International Conference on World Wide Web*, Pages: 507-517.
- Heydari, A., Tavakoli, M.A., Salim, N., Heydari, Z., (2015). Detection of Review Spam: A Survey, *Expert Systems with Applications*, 42:7, 3634-3642.
- Jindal, N., Liu, B. (2007). Review Spam Detection. *16th International Conference on World Wide Web*, Pages: 1189–1190.
- Jindal, N., Liu, B., (2008). Opinion Spam and Analysis, *WSDM '08 Proceedings of the 2008 International Conference on Web Search and Data Mining*, Pages 219-230.
- Jindal, N., Liu, B., Lim, E.P., (2010). Finding Unusual Review Patterns using Unexpected Rules, *CIKM '10 Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, Pages 1549-1552.
- Keshavaras, F., Waheed, A.A., Rachdi, B., Alhadj R., (2018), Review Spam Detection by Highlighting Potential Spammers and Diminishing Their Effect, *International Journal of E-Business Research*, 14(1): 23.
- Lam, S. K., Riedl, J. (2004). Shilling Recommender Systems for Fun and Profit, *13th International Conference on World Wide Web*, Pages: 393–402.
- Lim, E.P., Nguyen, V.A., Jindal, N., Liu, B., Lauw, H. W. (2010). Detecting Product Review Spammers using Rating Behaviors. *19th ACM International Conference on Information and Knowledge Management*, Pages: 939-948.
- Lin, Y., Zhu, T., Wang, X., Zhang, J., Zhou, A., (2014). Towards Online Review Spam Detection, *Proceedings of the 23rd International Conference on World Wide Web*, Pages: 341-342.
- Mukherjee, A., Liu, B., Wang, J., Glance, N., Jindal, N. (2011). Detecting Group Review Spam, *20th International Conference Companion on World Wide Web*, Pages: 93–94.
- Thet, T.T., Na, J.C., Khoo, C.S.G., (2010). Aspect-Based Sentiment Analysis of Movie Reviews on Discussion Boards, *Journal of Information Science*, 36:6, 823-848.
- Xie, S., Wang, G., Lin, S., Yu, P. S. (2012). Review Spam Detection via Time Series Pattern Discovery. *21st International Conference Companion on World Wide Web*, Pages: 635–636.

## Effectively Detection of Product Review Spammers

### Extended Abstract

*With increasing popularity of online shopping sites in recent years, internet users are purchasing more products over the Internet. Today, even market shopping can be done over the internet in many big cities in Turkey. It has been determined by many studies that users who shop over the Internet pay attention to reviews of other users about products, and often their buying trends change with the views of other users. Companies find sometimes volunteer and sometimes paid reviewers to create positive opinions about their products or to worsen good impression about competing products for the purpose of manipulate the reviewing system. Thus, they tried to change the opinions of users who would buy products, by writing reviews that are away from objectivity and sided. These reviews are too many to be rectified by human hands. Therefore, many approaches have been presented to detect spam reviews and spammers. With these approaches, the reliability of reviewing systems can be ensured by detecting spam reviews and spammers. Also, confidence in reviews can increase by providing readers to read not manipulated reviews.*

*In this study, a new approach is proposed by integrating four different methods that are proposed to find product review spammers. In the proposed method, 4 different review spammers finding approaches are put together. The method of finding review spammers for each approach is different from each other. The proposed method can thus find review spammers who exhibit rare or suspicious behavior or group activity. The first approach finds the reviewers with extreme ratings. In the second approach, spammers rate differently non-targeted products but they rate targeted products more differently. In the third approach, spammers who intensively comment*

*on a specific time interval are targeted. There is a high probability that the commentator who writes many comments on the positive or negative side during a given time period is a spammer. In the fourth and last approach, it is tried to determine spammers working in groups more than self-study.*

*The third part of the proposed architecture is called the mediator. Spammers in this part are sorted from large to small by using the proposed formula. According to this method, the highest number of commentators by a single approach is less likely to be fake than the others.*

*Some experiments have been made to test the performance of the proposed method. As data, comments written for Amazon products are used. The dataset used to obtain the order of precedence among approaches contain 1,689,188 comments, 382,176 items and 143,615 commentators.*

*In accordance with the purpose of the proposed system, spammers found only by two or more approaches are put into order. The first priority is the spammers found by the four approaches.*

*In the second experiment, a synthetic data set was created to test the proposed system, with 750 non-fake reviewers and 250 fake reviewers.*

*The last experiment is obtaining the accuracy value of the solutions with at least two approaches. Table 2 shows the accuracy of these approaches. As expected, the highest accuracy is the spammers found in all four approaches. The method with the lowest accuracy is the solutions obtained by approaches 2 and 4.*

*The method tested on Amazon product reviews exhibits more accurate results.*

**Keywords:** Review Spam, Review Spammer, Machine Learning, Classification